Evaluation of Children Cursive Handwritten Words for e-Education

Damien Simonnet, Nathalie Girard, Eric Anquetil, Mickaël Renault, Sébastien Thomas

Evaluation of Children Cursive Handwritten Words for e-Education

Damien Simonnet[a], Nathalie Girard[b], Eric Anquetil[a], Mickaël Renault[a], Sébastien Thomas[a]

[a] INSA Rennes, UMR IRISA, Campus de Beaulieu, F-35042 Rennes, France
[b] University Rennes 1, UMR IRISA, Campus de Beaulieu, F-35042 Rennes, France

Abstract

As part of an innovative e-education project, a digital workbook is being developed to help teach handwriting at school for children aged three to seven. The main objective of this project is to offer an advanced digital writing experience at school by using pen-based tablets. In this paper, an automatic qualitative analysis process of cursive handwriting words is presented. This approach is original because the goal is not to recognise the word that was handwritten by children (it is an explicit instruction) but to design a precise evaluation of the quality of his handwriting production to give them a real-time feedback. The presented method is based on a specific explicit elastic letter spotting segmentation able to deal with the imprecision of the handwriting of young children. This approach is suited to automatically and precisely highlight the difficulties encountered by children (adding or missing letters, incorrect shapes...). The validation of the proposed approach has been done on a dataset collected in French preschools and primary schools from 231 children. Beyond quantitative results, this paper reports the very positive impact of using this digital workbook that allows children to work independently with online and real-time feedbacks.

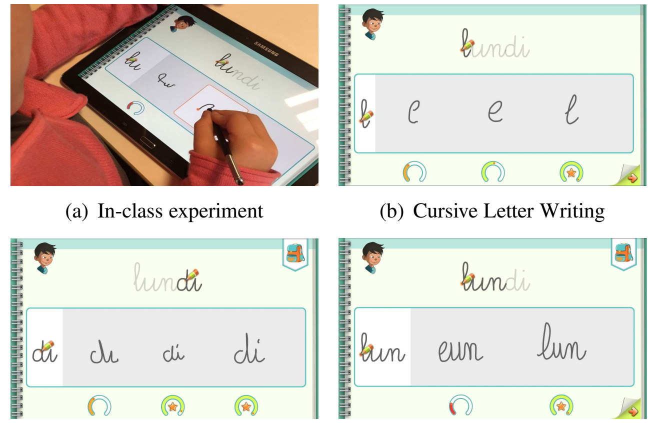
1. Introduction

Several studies show that digital devices can help students and teachers in a learning context. For example, Chickering and Gamson (1987) have extracted *seven principles to improve teaching*, like *Good practice uses active learning techniques*, *Good practice gives quick feedback*. Then, in Chickering and Stephen (1996), some ways to use new technologies are described to apply these principles in a learning context. This article focuses on handwriting learning by young children (three to seven years old) in primary schools. In an overview of the overall criteria used in scales for handwriting evaluation, Rosenblum et al. (2003) highlight that as a consensus about *which criteria constitute the critical components of handwriting readability, most researchers accept the criteria of size (height, width); slant; spacing (spaces between letters/words); the degree of line-straightness; shape (letter form and shape); and the general merit of the writing*. They also conclude that *computer based analysis are more accurate, sensitive, and reliable than the subjective analysis*, but they also observe that the current *practical applications are still limited*. In fact, these applications were *not yet capable of making a global decision as to the legibility of the writing product*. Since this study, digital devices have been improved. Especially, efficient pen-based tablets have been designed. This kind of tablet allows writing on screen with a sensation very similar to writing on paper, in particular because the user can hold the pen in the same way, and can also lay the palm of the hand on the screen as on paper. Moreover, Jolly et al. (2013) have presented a comparative study about the acquisition of handwriting between training with digital devices and paper. Results show a significant improvement of children trained on the digital device compared to children trained on paper, notably in terms of fluency (decreasing of the 'in-air' time, and stopping time). In this study, children writings are mainly analysed according to recorded time, velocity, etc. but letter height, width, form and shape are not considered. With pen-based tablet, various criteria such as shape, direction, order, pressure, fluidity can also be analysed. In this objective, a digital notebook was designed to help writing learning, see Fig. 1(a).

Moreover, the digital notebook gives immediate and personalised feedback to children. This prompt feedback helps them to be more autonomous. Another advantage of the notebook is the ability to adapt the educational progression to each child.

Since handwriting learning involves cognitive, kinesthetic and perceptual-motor components (Rosenblum et al., 2003; Djeziri et al., 2002; Plamondon et al., 2014), the notebook is based on several modules dedicated to different required skills, and is defined alongside pedagogical experts. The notebook is currently composed of six modules linked to different learning steps in primary schools: *Block Letter Writing*, *Digit Writing*, *Word Identification*, *Graphical Identification*, *Preparation to the Cursive Writing* and *Cursive Writing*. Moreover, in each module, special attention is paid to the feedback returned to children. Actually, prompt (*i.e.* real-time) and adaptive feedback are very important in learning processes (Kluger and DeNisi, 1996; Shute, 2008).

This paper focuses on the cursive writing module in which children work first on an individual letter (see Fig. 1(b)). Then, if the written production is correct, they work on bigrams, trigrams and words (see Fig. 1(c) and Fig. 1(d)). In this module (and for all modules), the objective is to evaluate the children's gesture (*i.e.* to identify if they use the right order and direction to write the letter, and if the shape of each letter is legible) in order to explain them how to correct and improve their writings. To do so, especially when more than one letter is written, **the evaluation implies to deal with the segmentation of a word in letters, and to identify matching, missing and added letters compared to the reference model.** In the literature, detection of a letter or a word as part of a sentence or a text has been extensively studied for handwriting recognition (Tappert et al., 1990; Plamondon and Srihari, 2000; Cheriet et al., 2007).

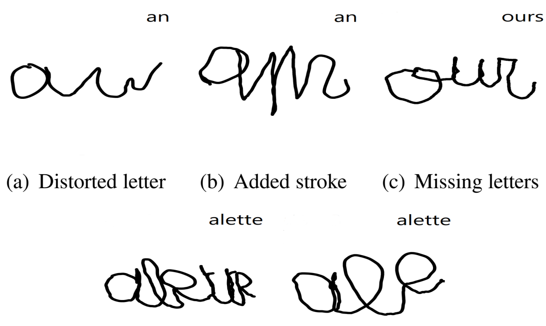


(a) In-class experiment (b) Cursive Letter Writing (c) Cursive bigrams Writing (d) Cursive trigrams Writing

Fig. 1. First in-class experiment of the e-education innovative project (a) with tablet touch-sensitive devices, and the cursive writing module scenario (b) (c) (d).

But this challenge is all the more complex in the context of the initial learning of the writing skills by children who have a very approximate handwriting. Indeed, they frequently add, distort, and even forget letters in the writing of a word. Moreover, children frequently combine several kinds of mistakes on the same word which makes the analysis task especially complex. As examples, Fig. 2(a) shows a distorted letter problem in the writing of the word *an* by a child, for the same word, Fig. 2(b) shows an addition of stroke that can induce confusion between *n* and *m*. Other common errors lie in missing letters as illustrated in Fig. 2(c) where the expected word is *ours* (*i.e.* bear). Fig. 2(d) shows a letter transformation, where the last *e* of the pseudo-word *alette* has been transformed in an *r*. Finally, Fig. 2(e), where the expected pseudo-word is *alette*, shows mistakes combination with missing letters and letter distortion.

Therefore, to tackle these issues, we propose to strengthen the common segmentations with verification steps based on specific analysis methods. The proposed workflow allows an evaluation of handwriting quality knowing the expected word, *i.e.* the teacher's instruction. In Section 2, the specific evaluation of children handwriting is presented. This approach has been tested on a real dataset and experimental results are pre-



(a) Distorted letter (b) Added stroke (c) Missing letters

(d) Transformed letter (e) Missing letters and distorted letter

Fig. 2. Children's approximate writings

sented in Section 3. Finally, Section 4 concludes this paper.

2. Evaluation of Children Handwritten Word Quality

In this section, a new approach is presented to evaluate the children's handwriting with regards to a targeted sequence, *i.e.* a grapheme or a word. The handwritten words are on-line signals captured on digital touch-screen devices with styluses. Since the main objective is to return corrective feedback to children, the system must detect correct letters, as well as incorrect writings, additional strokes and missing letters. To do so, as summarised in Fig. 3, the evaluation of the handwritten word quality consists of three main steps: at first, extraction of primary segmentation hypotheses, then extraction of letter hypotheses and finally word hypothesis extraction and evaluation.

As shown in Fig. 3, the evaluation is based on a standard workflow from an initial segmentation to word extraction with a recognition process. But as detailed in subsection 2.2, the standard recognition processes are inefficient on distorted writings like those of children. Thus, the letter hypotheses extraction step is reinforced to reduce error propagation by adding a verification step through a supervision (see the gray sub-block 2.2 in Fig. 3). The ranking of the word hypotheses is also redefined with a new scoring (gray sub-block 3.2 in Fig. 3). And, finally, we propose a new evaluation of the quality of the written word combining elastic matching and writing analysis score (gray sub-block 3.3 in Fig. 3).

As mentioned before, in our context the *expected word* is known, and refers to the sequence that should be written by children. Hereafter, the *expected word* is denoted $X = (x_i)_{i \in [\![1,m]\!]}$ where x_i represents a character, and the *analysed word* is denoted $Y = (y_i)_{i \in [\![1,n]\!]}$ where y_i represents a letter.

The rest of this section details each block in the workflow. Each subsection focuses on the proposed strengthening.

2.1. Extraction and Organisation of Segmentation Hypotheses

The first part of the evaluation of the children's handwriting consists in extracting *segmentation hypotheses* corresponding to sub-strokes of the original drawing: a letter, a sub-part of a letter, or more generally a sub-stroke that could be an expected letter. *Segmentation hypotheses* are build from the *primary segmentation* and organised in a graph.

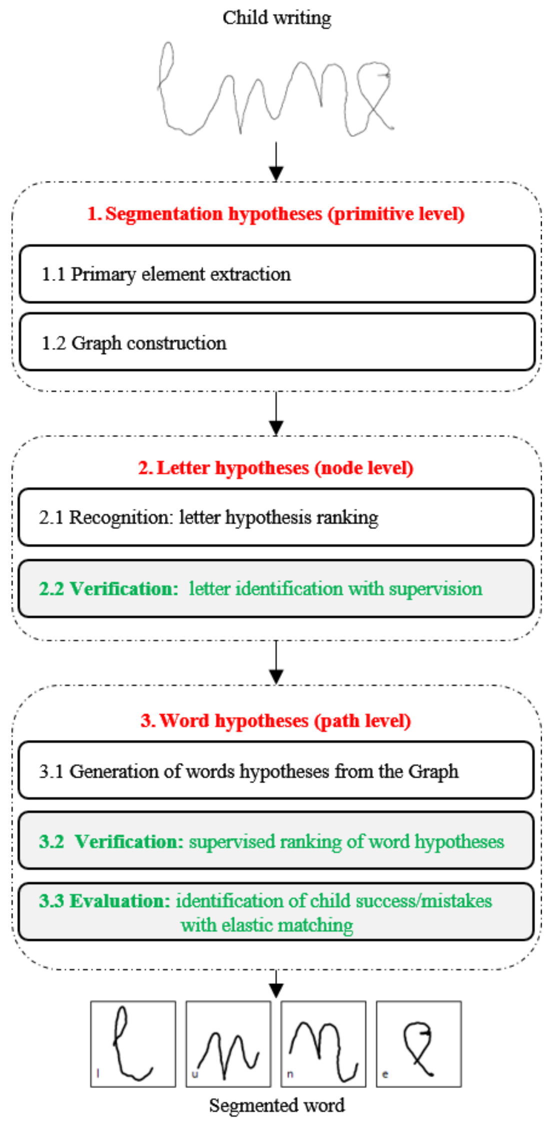
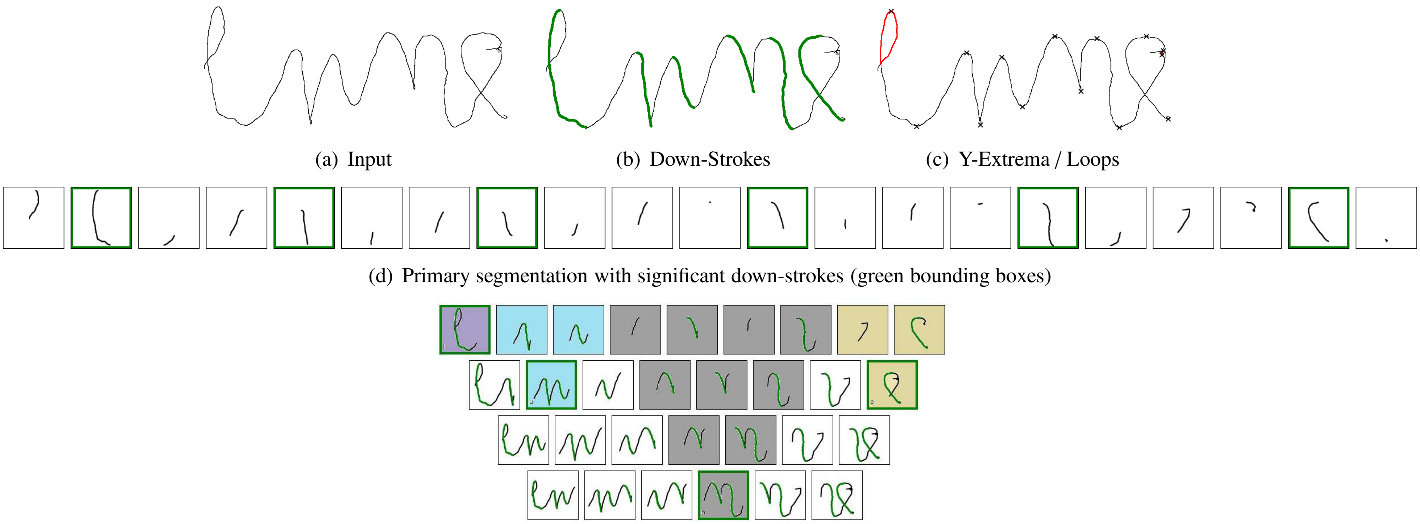


Fig. 3. Automatic Cursive Handwriting Evaluation overflow in which the expected word is 'lune' (moon).

2.1.1. Primary element extraction

The *primary segmentation* corresponds to a partition of the on-line handwritten signal based on stable parts of the handwriting: down-strokes (Anquetil and Lorette, 1997a). As an example, for the input word *lune* (*i.e.* moon) in Fig. 4(a), the primary segmentation is illustrated in Fig. 4(d). More precisely, this segmentation is spatially ordered along the x-axis, and has been built by considering significant down-strokes as indivisible (see green parts in Fig. 4(b) matching with green nodes in Fig. 4(d)). Cutting points are based on y-extrema (see each cross in Fig. 4(c)) and points at one third of the curvilinear distance for up-strokes non associated to loops (in red in Fig. 4(c)). This segmentation represents all possible cutting in the on-line signal around significant down-strokes.



(a) Input (b) Down-Strokes (c) Y-Extrema / Loops

(d) Primary segmentation with significant down-strokes (green bounding boxes)

(e) Segmentation Graph (four levels)

Fig. 4. Segmentation Graph (e) and Primary Segmentation (d) of an input gesture *lune* (a) based on loops and y-extrema (c), and down-strokes (b). Loops and down-strokes are represented by red and green drawing, crosses represent y-extrema.

2.1.2. *Graph construction*

Similarly to (Anquetil and Lorette, 1997b), a segmentation graph (Fig. 4(e)) is used to represent all partitions of the input handwritten signal. This graph contains nodes corresponding to *segmentation hypotheses* that are valid for the cursive Latin alphabet when there are between one and three significant down-strokes. The first level of the *segmentation graph* results from the *primary segmentation* in which each down-stroke has been merged with adjacent up-strokes resulting in the level where only one down-stroke is used. The next levels of the graph are built by creating *segmentation hypotheses* corresponding to the merger of all consecutive hypotheses of the previous level. This process through levels is depicted on Fig. 4(e) by node colourisation, where the lowest gray node is the merger of the two gray nodes of the previous level, which are themselves the merger of the gray nodes of the previous level and so on.

After the graph construction, the *identification process* of *letter hypotheses* is conducted in each graph node. The next section presents this step.

2.2. *Identification of Word Letter Hypotheses*

The identification process consists in finding the most relevant *letter hypotheses* associated to a *segmentation hypothesis* knowing the *expected word*. This process combines two complementary steps: *prediction* and *verification*. The former, associated to a recognition process, predicts the most relevant symbols of the alphabet for a given *segmentation hypothesis*. The latter checks the correctness of predicted hypotheses through two analysis processes: local and global.

More precisely, for each *segmentation hypothesis* the recognition process gives a recognition confidence for each letter of the alphabet. On the contrary, the analysis process takes as input a *segmentation hypothesis* with a letter hypothesis, and returns an analysis confidence.

(a) (b) (c)

Fig. 5. Letter hypotheses computation from the last node of the second level of the graph in Fig. 4(e) with recognition scores and ranking in (a), analysis scores and ranking in (b), and global letter hypothesis selection in (c)

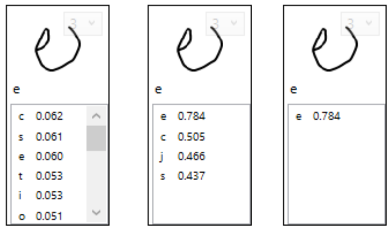
2.2.1. *Local Prediction of Letter Hypotheses*

The local prediction process consists in finding the most relevant *letter hypotheses* associated to a *segmentation hypothesis* (*i.e.* a graph node) based on the recognition confidences. In a standard approach, the best letter based on recognition confidence would be used as correct hypotheses. However, as illustrated in Fig. 5(a) where the expected letter is an 'e', the recognition process (here an *EVOLVE* classifier (Almaksour and Anquetil, 2011, 2013), see Section 3.2 for details) recognises as best hypothesis an 'l' due to the degraded writing of child. If this letter is considered, the feedback to child will be inadequate.

In our context, the expected word is known and can be used to select the most relevant hypotheses in a verification step. Moreover, this verification step should guide the letter hypotheses extraction to allow useful feedback. This original verification step is explained in the following section.

2.2.2. *Local and Global Verification of Letter Hypotheses*

The verification step is fundamental to analyse finely handwriting that allows to understand degraded children's handwrit-



(a) (b) (c)

Fig. 6. Letter hypotheses computation from letter 'e' in word 'je' ('i') with recognition scores and ranking in (a), analysis scores and ranking in (b), and global letter hypothesis selection in (c)

ing.

The local verification process consists first in selecting the most relevant hypotheses with regards to the expected sequence X and then in checking their correctness with an analysis process. As an example, in Fig. 5(b), the expected letter is an 'e', the analysis score is computed based on this knowledge and letter hypotheses are re-ranked according to this score (here the analysis score is computed based a confidence-based classifier (Simonnet et al., 2017), see Section 3.2 for details). This re-ranking takes into account the letter confusion children can have in a learning context. It also allows to pull correct letter up in order to give useful feedback. Therefore, hypotheses are selected from two sources: the *expected letters* (n_h^a best) corresponding to the context and the recognition process (n_h^r best) to handle confusion. Then, an analysis process is used to evaluate the quality and the relevance of the selected hypotheses. If the analysis score is smaller than the correct analysis rate θ_a, the hypothesis is discarded as the writing is too degraded to be considered as a valid hypothesis.

The global verification process considers all *letter hypotheses* resulting from the local verification process and select the most adapted to the expected sequence X knowing that some parts of the children writing may not be legible. The first part of the selection keeps the n_{best} *letter hypotheses* of each character in the sequence X. As an example, in Fig. 5(c) and Fig. 6(c), 'e' is expected in the word and its analysis score is high compared to other 'e' letter hypotheses in the graph, therefore this letter hypothesis is kept in this node. However, due to non legible parts of children writing, the previous process may remove all *letter hypotheses* in a graph node. A second process is therefore used to add *letter hypotheses* with the best recognition score, which implies that all nodes of the graph have at least one *letter hypothesis*. Thus, this last step allows to identify unexpected letters written by children, *i.e.* letters which are not in the teacher's instructions.

2.3. *Selection of the Best Word Segmentation Path*

The last part of the handwriting evaluation consists in extracting the most probable partition of the handwriting input signal with regards to the *expected word* called *best word segmentation path*. The previous letter hypotheses combined with the segmentation path of the graph allow to build a set of *word segmentation path (WSP)*. To find *the best WSP*, an analysis score is computed for each hypothesis, and finally this score is used with an elastic matching method to extract the best analysis path.

In the rest of the section, first the notion of *segmentation path* is presented, followed by the process of selection of WSP hypotheses. Finally, the best WSP hypothesis is selected with an elastic matching method.

2.3.1. *Segmentation Path*

A *segmentation path* is a path in the graph corresponding to a partition of the handwriting's on-line signal. More formally, let $P_n^a = (S^i_{l_i,k_i})_{i\in[\![1,n]\!]}$ be a partition of the handwriting's on-line signal where $S^i_{l_i,k_i}$ corresponds to a substroke of the input signal associated to the level l_i and position k_i of the segmentation graph, and n is the size of the *segmentation path*.

2.3.2. *Selection of Word Segmentation Path Hypotheses*

The first step in the extraction of the *best word segmentation path* builds a set of hypotheses reflecting children writing based on the *segmentation path* of the graph and letter hypotheses selected. The selection process based on an analysis path score s_{ap} detects hypotheses corresponding to various scenarios of the children writing (noisy strokes, confused or forgotten letters).

The analysis path score s_{ap} combines an analysis score s_a representing the correctness of each segmented element with consecutive coherence scores (n-gram and inter-letter matching) as follows:

$$s_{ap}(P_n^a) = \frac{1}{2}s_a(P_n^a) + \frac{1}{4}s_t^{ng}(P_n^a) + \frac{1}{4}s_t^{im}(P_n^a) \tag{1}$$

The first coherence score s_t^{ng} corresponds to a coherence in relation with the *expected* word by penalising gradually bi-gram different from the *expected* word as expressed in equation (2).

$$s_t^{ng}t(i, i-1) = \begin{cases} 1.0 & \text{if} & (i, i-1) \text{ is a bi-gram of } X \\ 0.7 & \text{if} & i \text{ or } i-1 \text{ is a letter of } X \\ 0.4 & \text{otherwise} & \end{cases} \tag{2}$$

The second coherence score s_t^{im} is related to the spatial coherence of letters as in (Anquetil and Lorette, 1997b).

Coherence scores are computed as the conjunction of transition scores $s_{t(i,i-1)}$ as presented in equation (3).

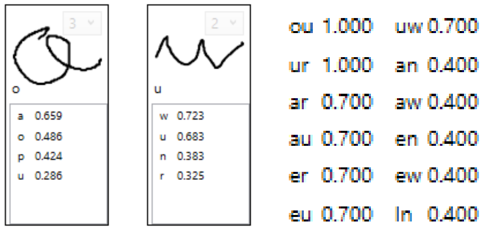
$$s_t(P_n^a) = \sqrt[n-1]{\Pi_{i=2}^n s_{t(i,i-1)}} \tag{3}$$

Finally, the analysis score s_a represents the average correctness of each element of the partition. It is defined mathematically by equation (4) where $s_a(i)$ is the analysis score of the $i-th$ path element.

$$s_a(P_n^a) = \sqrt[n]{\Pi_{i=1}^n s_a(S^i_{l_i,k_i})} \tag{4}$$

The N_{WSPH} best paths according to the analysis path score s_{ap} are kept as candidates to the selection of the best hypothesis.

As shown in Fig. 7(a) and Fig. 7(b), the analysis scores for the first letters of the word would suggest an 'a' followed by a 'w' but the information about the expected word 'ours' (bear) will increase the score of the bi-gram 'o-u' over the score of the bi-gram 'a-w' (Fig. 7(c)).



(a) (b) (c)

Fig. 7. Segmentation path hypotheses computation from word 'ours' (bear) with first letter analysis scores and ranking in (a), second letter analysis scores and ranking in (b), and some bigram coherence scores s_t^{ng} in (c)

2.3.3. Selection of the Best Hypothesis

A customised elastic matching method is used to match the *expected word X* asked by teacher with the *analysed word Y* that is associated to an analysis path P_n^a. This step makes it possible to identify correct letters, substitutions, missing and additional letters. More precisely, a Damerau Levenshtein edit distance (Damerau, 1964) is computed between X and Y, with optimised substitution costs learnt from the analyser. Insertion and deletion costs are proportional to the number of down-strokes. The substitution cost is composed of two parts: the first is learnt from the analyser and the second is a penalisation when the number of significant down-strokes between the two letters to substitute is different. The learnt substitution cost corresponds to an analysis confusion score defined by equation (5) where $s_a^l(k)$ represents the analysis score of a symbol k as a symbol l, and K, L are respectively the training set for the symbol k and l.

$$c^l_{substition}(k,l) = 1 - \frac{\sum_{k \in K} s_a^l(k)}{\sum_{l \in L} s_a^l(l)} \tag{5}$$

The final substitution cost is given by equation (6) where $\alpha_{k,l}$ is the penalisation factor defined in equation (7) and depends on the absolute difference of significant down-strokes $d^a_{k,l}$.

$$c_{substition}(k,l) = \alpha_{k,l} \times c^l_{substition}(k,l) \tag{6}$$

$$\alpha_{k,l} = \begin{cases} 1.0 & \text{if} \quad d^a_{k,l} = 0 \\ 0.5 + d_{k,l} & \text{else} \end{cases} \tag{7}$$

Finally, the best analysis path is chosen as the analysis path hypothesis with the minimum edit distance.

3. Experiments and Results

This section presents experimental results of the analysis segmentation algorithm. First, the testing dataset is introduced followed by the recognition and analysis processes used for these experiments. Then, the evaluation protocol and results are explained.

3.1. Datasets

The testing dataset was collected during in-class experiments in a context of a user-centered design approach that uses feedback of children and teachers to improve the application. More specifically, during digital workshops of twenty minutes with six tablets children were drawing sequences and letters.

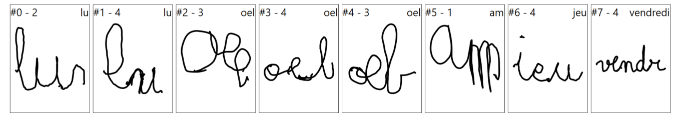
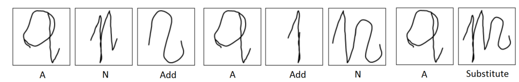


Fig. 8. Sample words in the testing dataset written by children with additional strokes (#0, #5), missing letters (#7), substitution (#4, #6) and deformed words (#1, #2, #3). The word that children should write is displayed on the top right corner.

(a) *an* (b) Correct letters



(c) The three Correct Word Segmentations

Fig. 9. Multiple word segmentations of an input handwriting signal (a) corresponding to an *expected* sequence *an*. (b) represents the correct letters with regards to the expected sequence. The three correct segmentations (c) of (a) correspond to sequences with addition and substitution of letters.

This dataset, composed of 1012 pseudo-words from two to eight letters, was collected in French preschools from 231 children from four to five years old. The list of all the sequences used in this dataset is ab, alette, am, an, ap, apin, di, es, qu, je, jeu, jeudi (= Thursday), ji, lu, lune (= moon), lundi (= Monday), mardi (= Tuesday), oel, onne, ours (= bear), oyeux, ph, rs, se, ux, ve, ven, vendredi (= Friday), samedi (= Saturday). In Fig. 8, there are some samples of this dataset containing writing such as additional strokes, missing letters, substitutions and deformed words. A segmentation ground truth is created semi-automatically for each sequence that can contain missing or additional strokes. First of all, segmented areas corresponding to letters are manually annotated, then an automatic process generates the segmented partition corresponding of the handwritten signal input including notably parts that do not correspond to segmented letters.

3.2. Recognition and Analysis Processes

For these experiments, the recognition process is based on the discriminative properties of a classifier that rank hypotheses of the alphabet. In this work, an *EVOLVE* classifier (Almaksour and Anquetil, 2011, 2013) is trained with *HBF49* (Delaye and Anquetil, 2013) features on the letters of the Latin alphabet. More precisely, *HBF49* is a generic set of features designed for handwriting symbols recognition. It is composed of *dynamic* features that depend on the writing process (*e.g.* starting and ending positions, proportion of down-strokes trajectory, angle of the initial vector, inflexions), and *visual* features that focus on the appearance of the written results (*e.g.* 2D histogram of point, k-perpendicularity, k-angle).

The analysis process uses cluster models built with *HBF49* features for each letter of the alphabet represented as a mean vector and a covariance matrix (Almaksour and Anquetil, 2011, 2013). More specifically, a confidence-based classifier (Simonnet et al., 2017) based on discriminative and generative (cluster) models with *HBF49* features is used to compute a normalised analysis score. This score is based on the calibration of the Mahalanobis distance to have a confidence score.

3.3. Evaluation Protocol

The analysis segmentation algorithm presented in this paper makes it possible to identify correct and incorrect parts of the drawing for a given input handwritten signal by returning an analysis path $P^a_n = (S^i_{l_i,k_i})_{i\in[\![1,n]\!]}$. The evaluation process quantifies the coherence of the segmentation with regards to the *expected* sequence asked.

It is noted that a written sequence can have several correct segmented partitions (*e.g.* an *expected* sequence *'an'* written as a *'am'* has three valid partitions as depicted in Fig. 9). Let $G_m = (S^i_g)_{i\in[\![1,m]\!]}$ be a ground truth partition of size m, and $\mathcal{G}$ the set of ground truth partitions associated to an input handwriting signal.

First, the evaluation process selects the closest ground truth to the analysis path by maximising the *matching score* defined in equation (8) where $d^{(i,j)}_{tp}$, $d^{(i,j)}_{fp}$, $d^{(i,j)}_{fn}$ correspond respectively to the curvilinear distance of true positive, false positive and false negative strokes between the substrokes S^i_g and S^j. The association between substrokes of a ground truth partition and an analysis path is done by first selecting the pair with the best matching score.

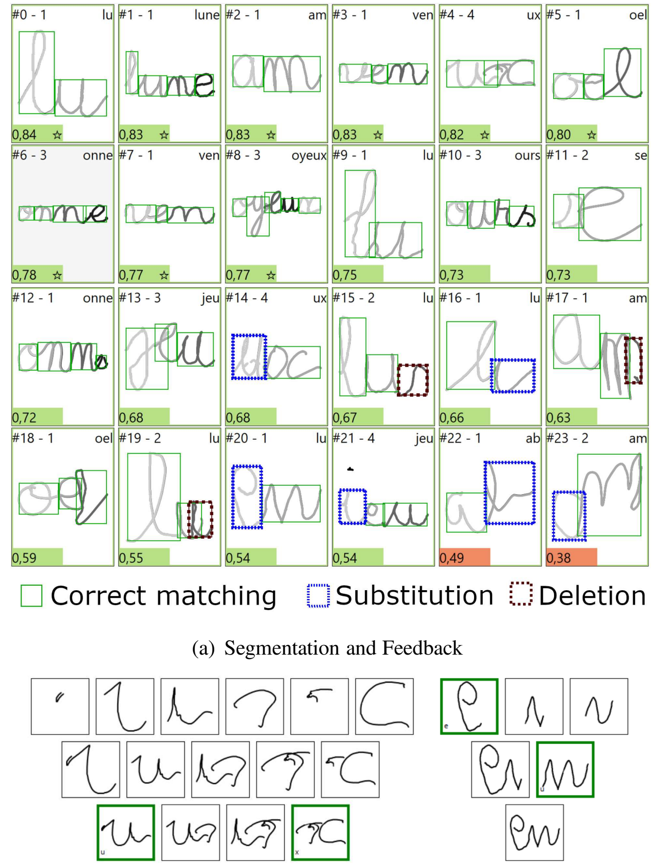
$$m(S^i_g, S^j) = \frac{d^{(i,j)}_{tp}}{d^{(i,j)}_{tp} + d^{(i,j)}_{fp} + d^{(i,j)}_{fn}} \quad (8)$$

From this process, two evaluation metrics are defined: the *letter average matching score* and the *correct word segmentation ratio*, that correspond respectively to the average of all the *matching scores*, and to the ratio of sequences correctly segmented, *i.e.* sequences having a *matching score* for all letters greater the τ_m threshold.

3.4. Results

This section presents quantitative and qualitative results of the analysis segmentation algorithm. In each node, the number of analysis and recognition hypothesis are respectively set up to $n^a_h = 2$ and $n^r_h = 2$. These two values have been chosen empirically to reduce the number of word path hypotheses. The correct analysis rate θ_a to reject incorrect analysis hypothesis is equal to 0.25 because in Simonnet et al. (2017) a score lower corresponds to an incorrect writing. The number of hypotheses kept corresponding to each letter of the sequence to analyse is empirically fixed to $n_{best} = 3$ to reduce the number of word path hypotheses. The maximum number of word segmentation path hypotheses kept for the elastic matching is empirically fixed to $N_{WSPH} = 100$. The matching threshold for the evaluation is $\tau_m = 0.7$.

The quantitative results on the testing dataset correspond to a *correct segmentation ratio* of 0.90 in which a *letter average matching score* is of 0.95.



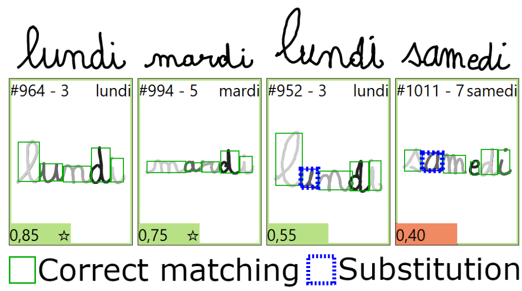
(a) Segmentation and Feedback

(b) ux (#4)

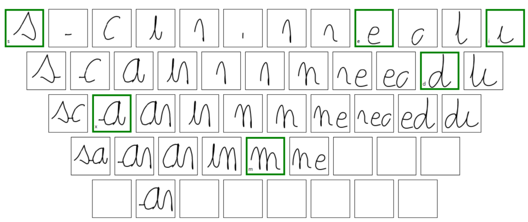
(c) lu (#20)

Fig. 10. Sequences Qualitative Results. In (a), the segmentation of the handwriting is represented with different greyscales. Green, blue and red bounding boxes correspond respectively to a letter with a correct matching, a substitution and a deletion. (b) and (c) are the segmentation graph of the sample #4 and #19. The score corresponds to the recognition path score defined in equation (1).

In Fig. 10 and Fig. 11 qualitative results are presented respectively for sequences and words where the segmentation of the handwriting is represented with different greyscales. Green, blue and red bounding boxes correspond respectively to a letter with a correct matching, a substitution and a deletion. In Fig. 10, results show that the presented method is able to segment children handwriting, and to identify additional strokes (#15 and #17) and some substitution (#20). However, although the segmentation is correct, some detected substitutions are incorrect (#14, #16) due to deformed letters. Results on words presented in Fig. 11 show the ability of the presented method to deal with longer sequences. Fig. 11(b) presents the segmentation graph of the word 'samedi' with the analysis path detected. Errors in the segmentation presented in Fig. 12 are mainly related to low handwriting quality of children (#0, #1, #2). In #3, the segmentation of the letter *u* is false because the additional part after the *u* is not detected as a segmentation hypothesis since it is composed of up-strokes only.

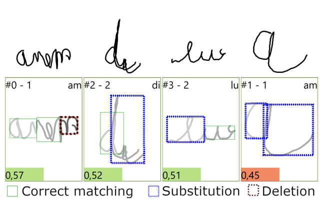


(a) Segmentation and Feedback

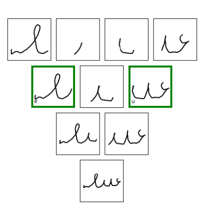


(b) samedi (#1011)

Fig. 11. Words Qualitative Results



(a) Segmentation and Feedback



(b) lu (#3)

Fig. 12. Errors of segmentation on children's handwriting (a), the word that children should write is indicated on the top left corner. (b) is the segmentation graph of #3.

4. Conclusion

This paper presents an original analysis method for handwriting quality evaluation, that is able to deal with children cursive handwriting. The proposed approach was defined to be well adapted for degraded handwriting analysis. Indeed, a common recognition method is reinforced with specific verification steps, and useful information extracted from a customised elastic letter spotting is used to give relevant feedback to children. This proposal allows a detailed evaluation of children's handwriting (added strokes, missing strokes and incorrect writing). The proposed analysis enables the definition of real-time feedback needed in learning systems, especially to make children more autonomous in their learning. This method gives very relevant results and is currently tested on a large scale (in 40 primary schools with 1000 children).

Acknowledgments

This work takes place in the context of a three year research project founded by the French government as part of innovative projects (BPI): *IntuiScript* (http://intuiscript.com/). This work has also been supported by the LabCom 'ScriptAndLab' founded by the ANR (ANR-16-LVC2-0008-01). The authors are grateful to their industrial partner (Script&Go), educational experts of Rennes academy, ESPE, Brittany region and LOUSTIC laboratory for their collaboration.

References

Almaksour, A., Anquetil, E., 2011. Improving premise structure in evolving takagi–sugeno neuro-fuzzy classifiers. Evolving Systems 2, 25–33.

Almaksour, A., Anquetil, E., 2013. ILClass: Error-driven antecedent learning for evolving takagi-sugeno classification systems. Applied Soft Computing 19, 419–429.

Anquetil, E., Lorette, G., 1997a. On line handwriting character recognition system based on hierarchical qualitative fuzzy modeling. Handwriting Recognition , 109–116.

Anquetil, E., Lorette, G., 1997b. Perceptual model of handwriting drawing. application to the handwriting segmentation problem, in: Proceedings of the Fourth International Conference on Document Analysis and Recognition, pp. 112–117. doi:`10.1109/ICDAR.1997.619824`.

Cheriet, M., Kharma, N., Liu, C.L., Suen, C., 2007. Character Recognition Systems: A Guide for Students and Practitioners. John Wiley & Sons.

Chickering, A.W., Gamson, Z.F., 1987. Seven principles for good practice in undergraduate education. AAHE Bulletin 39(7), 3–7.

Chickering, A.W., Stephen, C.E., 1996. Implementing the seven principles: Technology as Lever. AAHE Bulletin 49(2), 3–6.

Damerau, F.J., 1964. A technique for computer detection and correction of spelling errors. Commun. ACM 7, 171–176.

Delaye, A., Anquetil, E., 2013. HBF49 feature set: A first unified baseline for online symbol recognition. Pattern Recognition 46, 117–130.

Djeziri, S., Guerfali, W., Plamondon, R., Robert, J., 2002. Learning handwriting with pen-based systems: computational issues. Pattern Recognition 35, 1049 – 1057. URL: `http://www.sciencedirect.com/science/article/pii/S0031320301000930`, doi:`http://dx.doi.org/10.1016/S0031-3203(01)00093-0`. handwriting Processing and Applications.

Jolly, C., Palluel-Germain, R., Gentaz, E., 2013. Evaluation of a tactile training for handwriting acquisition in french kindergarten children: A pilot study. Kindergartens: Teaching methods, expectations and current challenges , 161–176.

Kluger, A.N., DeNisi, A., 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin 119, 254–284.

Plamondon, R., O'Reilly, C., Galbally, J., Almaksour, A., Anquetil, É., 2014. Recent developments in the study of rapid human movements with the kinematic theory: Applications to handwriting and signature synthesis. Pattern Recognition Letters 35, 225–235.

Plamondon, R., Srihari, S.N., 2000. Online and off-line handwriting recognition: a comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 63–84. doi:10.1109/34.824821.

Rosenblum, S., Weiss, P.L., Parush, S., 2003. Product and process evaluation of handwriting difficulties. Educational Psychology Review 15, 41–81.

Shute, V.J., 2008. Focus on formative feedback. Review of Educational Research 78, 153–189.

Simonnet, D., Anquetil, E., Bouillon, M., 2017. Multi-criteria handwriting quality analysis with online fuzzy models. Pattern Recognition , –URL: http://www.sciencedirect.com/science/article/pii/S0031320317301474, doi:http://doi.org/10.1016/j.patcog.2017.04.003.

Tappert, C.C., Suen, C.Y., Wakahara, T., 1990. The state of the art in online handwriting recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 787–808. doi:10.1109/34.57669.